GEORGIA INSTITUTE OF TECHNOLOGY
OFFICE OF CONTRACT ADMINISTRATION
SPONSORED PROJECT INITIATION

Date: June 23,1976

Project Title: AFAPL Senior Investigator Program

Project No: E-21-682

Project Director: Dr. R. P. Webb

Sponsor: Southeastern Center for Electrical Engineering Education, Inc.
(Subcontract under AFAPL Prime No.F33615-76-C-2050)

Agreement Period: From February 1, 1976 Until February 1, 1979

Type Agreement: Subcontract No. SCEEE-AFAPL/SIP-1-76

Amount: $90,000 (Partially funded at $13,538)
1,354 GIT (E-21-344)
$14,892 TOTAL

Reports Required: Quarterly Progress Reports; Interim (Final) Report

Sponsor Contact Person (s):

Technical Matters

Contractual Matters
(thru OCA)

Southeastern Center for Electrical Engineering Education, Inc.
Room 202 D, Samford Hall
Auburn University
Auburn, Alabama 36830

Defense Priority Rating: DO-C9 under DMS Reg.1.

Assigned to: School of Electrical Engineering (School/Laboratory)

COPIES TO:

- Project Director
- Division Chief (EES)
- School/Laboratory Director
- Dean/Director—EES
- Accounting Office
- Procurement Office
- Security Coordinator (OCA)
- Reports Coordinator (OCA)
- Library, Technical Reports Section
- Office of Computing Services
- Director, Physical Plant
- EES Information Office
- Project File (OCA)
- Project Code (GTRI)
- Other

GEORGIA INSTITUTE OF TECHNOLOGY
OFFICE OF CONTRACT ADMINISTRATION

SPONSORED PROJECT TERMINATION

Date: February 18, 1980

Project Title: AFAPL Senior Investigator Program

Project No: E-21-682

Project Director: Dr. R. P. Webb

Sponsor: Southeastern Center for Electrical Engineering Education, Inc.

Effective Termination Date: 9/30/77

Clearance of Accounting Charges: 9/30/77

Grant/Contract Closeout Actions Remaining:

None

TERMINATED

_ Final Invoice and Closing Documents
_ Final Fiscal Report
_ Final Report of Inventions
_ Govt. Property Inventory & Related Certificate
_ Classified Material Certificate
_ Other _____

Assigned to: Electrical Engineering (School/Laboratory)

COPIES TO:

Project Director
Division Chief (EES)
School/Laboratory Director
Dean/Director–EES
Accounting Office
Procurement Office
Security Coordinator (OCA)
✓Reports Coordinator (OCA)
OCA Research Property Coordinator
Library, Technical Reports Section
EES Information Office
Project File (OCA)
Project Code (GTRI)
Other _____

CA–4 (1/79)

Summary Report

An Analysis of System Identification Techniques Suitable for Synchornous Machine Modeling

Prepared for:

AFAPL Senior Investigator Program
Contract No. F33615-74-C-2039

by

Mr. C. W. Brice
Dr. R. P. Webb
Dr. M. E. Womble

Electric Power Laboratory
School of Electrical Engineering
Georgia Institute of Technology

I. INTRODUCTION

In order to simulate a synchronous electrical generator under severly unbalanced loads, using state variable representation, it is imperative that the parameters of the differential equation be accurately estimated. Standard IEEE machinery tests (10) lead to gross inaccuracies in determining these parameters (9), primarily due to the numerous intermediate calculations required. The purpose of this report is to present the results of a comparative analysis of several techniques of directly estimating the parameters via system identification methods which have had extensive application in modern control theory (3,4,5,6). Three estimation techniques, namely the Bayesian-state augmentation method, the method of maximum likelihood and the method of weighted least squares, are compared and contrasted with reference to their applicability to the particular problems of machine parameter estimation. A brief review of each method is followed by a summary of the advantages and disadvantages of each method. Finally, a numerical example applying the method of weighted least squares to a simplified representation of the system is presented and discussed, and a method is proposed which alleviates numerical problems associated with radically different time constants.

II. Machine Equations

The equations of motion of a synchronous machine in d-q-0 coordinates (Park's equations) (9, 11, 13) can be expressed in state variable form as follows:

$$\dot{x} = \begin{bmatrix} \dot{\lambda}_d \\ \dot{\lambda}_f \\ \dot{\lambda}_{kd} \\ \dot{\lambda}_q \\ \dot{\lambda}_{kq} \\ \dot{\lambda}_0 \end{bmatrix} = \omega \begin{bmatrix} \frac{-R_d}{x''_d} & \frac{R_d}{x''_{df}} & \frac{R_d}{x''_{kdd}} & 1 & 0 & 0 \\ \frac{R_f}{x''_{df}} & \frac{-R_f}{x''_f} & \frac{R_f}{x''_{kdf}} & 0 & 0 & 0 \\ \frac{R_{kd}}{x''_{kdd}} & \frac{R_{kd}}{x''_{kdf}} & \frac{-R_{kd}}{x''_{kd}} & 0 & 0 & 0 \\ -1 & 0 & 0 & \frac{-R_q}{x''_q} & \frac{R_q}{x''_{kqq}} & 0 \\ 0 & 0 & 0 & \frac{R_{kq}}{x''_{kqq}} & \frac{-R_{kq}}{x''_{kq}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{-R_0}{x_0} \end{bmatrix} \begin{bmatrix} \lambda_d \\ \lambda_f \\ \lambda_{kd} \\ \lambda_q \\ \lambda_{kq} \\ \lambda_0 \end{bmatrix}$$

$$+ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_d \\ V_f \\ V_q \\ V_0 \end{bmatrix} = AX + BU$$

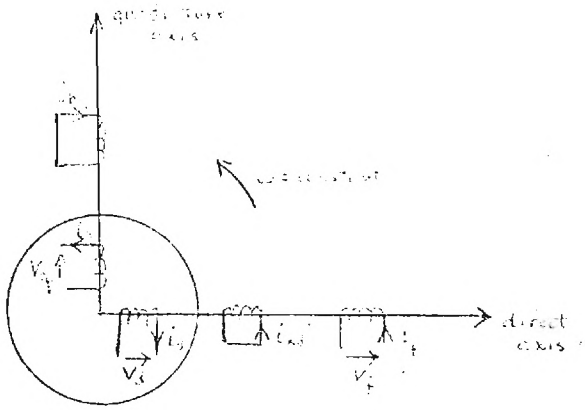
$$i = \begin{bmatrix} -i_d \\ i_f \\ i_{kd} \\ -i_q \\ i_{kq} \\ -i_0 \end{bmatrix} = \omega \begin{bmatrix} \frac{1}{x''_d} & \frac{-1}{x''_{df}} & \frac{-1}{x''_{kdd}} & 0 & 0 & 0 \\ \frac{-1}{x''_{df}} & \frac{1}{x''_f} & \frac{-1}{x''_{kdf}} & 0 & 0 & 0 \\ \frac{-1}{x''_{kdd}} & \frac{-1}{x''_{kdf}} & \frac{1}{x''_{kd}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{x''_q} & \frac{-1}{x''_{kqq}} & 0 \\ 0 & 0 & 0 & \frac{-1}{x''_{kqq}} & \frac{1}{x''_{kq}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{x_o} \end{bmatrix} \begin{bmatrix} \lambda_d \\ \lambda_f \\ \lambda_{kd} \\ \lambda_q \\ \lambda_{kq} \\ \lambda_0 \end{bmatrix} = Y_{sc}\lambda$$

$$z = \begin{bmatrix} i_a \\ i_b \\ i_c \\ i_f \end{bmatrix} = \sqrt{2/3} \begin{bmatrix} -\cos(\omega t) & 0 & 0 & +\sin(\omega t) & 0 & \frac{-1}{\sqrt{2}} \\ -\cos(\omega t-120^\circ) & 0 & 0 & +\sin(\omega t-120^\circ) & 0 & \frac{-1}{\sqrt{2}} \\ -\cos(\omega t+120^\circ) & 0 & 0 & +\sin(\omega t+120^\circ) & 0 & \frac{-1}{\sqrt{2}} \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -i_d \\ i_f \\ i_{kd} \\ -i_q \\ i_{kq} \\ -i_0 \end{bmatrix} = Ti$$

thus,

$$z = TY^x = Cx$$

where λ represents flux linkages, i represents currents, and the matrix Y_{sc} is just the short-circuit admittance matrix of multiport circuit theory. A schematic diagram of the machine illustrating the sign conventions, which are the standard generator notations, follows,



The above equations assume that the speed remains constant and neglect non-linear effects such as saturation. The equations have been normalized and are expressed in the so called per unit system (8, 9, 12) that is the reactances and resistances are dimensionless ratios of the actual values in ohms to a base impedance determined from machine ratings.

If the load on the generator is assumed to be an unbalanced resistive one, i.e. R_a, R_b, and R_c, then the voltages v_a, v_b, and v_c are related to the currents i_a, i_b, and i_c by

$$\begin{bmatrix} v_a \\ v_b \\ v_c \end{bmatrix} = \begin{bmatrix} R_a & 0 & 0 \\ 0 & R_b & 0 \\ 0 & 0 & R_c \end{bmatrix} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} .$$

Carrying out the time varying transformations relating these to the d-q-0 coordinate gives

$$\begin{bmatrix} v_d \\ v_q \\ v_0 \end{bmatrix} = \begin{bmatrix} R_{L0} + R_{L1}(t) & R_{L2}(t) & \sqrt{2}\, R_{L3}(t) \\ R_{L2}(t) & R_{L0} - R_{L1}(t) & \sqrt{2}\, R_{L4}(t) \\ \sqrt{2}\, R_{L3}(t) & \sqrt{2}\, R_{L4}(t) & R_{L0} \end{bmatrix} \begin{bmatrix} i_d \\ i_q \\ i_0 \end{bmatrix},$$

where

$$R_{L0} \triangleq \frac{1}{3}(R_a + R_b + R_c)$$

$$R_{L1}(t) \triangleq \frac{1}{3}[R_a \cos 2\omega t + R_b \cos 2(\omega t - 120^\circ) + R_c \cos 2(\omega t + 120^\circ)]$$

$$R_{L2}(t) \triangleq -\frac{1}{3}[R_a \sin 2\omega t + R_b \sin 2(\omega t - 120^\circ) + R_c \sin 2(\omega t + 120^\circ)]$$

$$R_{L3}(t) \triangleq \frac{1}{3}[R_a \cos \omega t + R_b \cos(\omega t - 120^\circ) + R_c \cos(\omega t + 120^\circ)]$$

and $$R_{L4}(t) \triangleq -\frac{1}{3}[R_a \sin \omega t + R_b \sin(\omega t - 120^\circ) + R_c \sin(\omega t + 120^\circ)].$$

After converting the d-q-0 currents to flux linkages (by multiplying by a matrix like Y_{sc}) and bringing these terms into the matrix A, the equations can be written as

$$\dot{x} = Fx + Gu \quad \text{and} \quad z = Hx$$

Now x and z remain the same, but $u = v_f$ and

$$F=\omega \begin{bmatrix} \frac{-(R_D+R_{L1}(t))}{x''_d} & \frac{R_D+R_{L1}(t)}{x''_{df}} & \frac{R_D+R_{L1}(t)}{x''_{kdd}} & 1-\frac{R_{L2}(t)}{x''_q} & \frac{R_{L2}(t)}{x''_{kqq}} & -\sqrt{2}\,\frac{R_{L3}(t)}{x_0} \\ \frac{R_f}{x''_{df}} & \frac{-R_f}{x''_f} & \frac{R_f}{x''_{kdf}} & 0 & 0 & 0 \\ \frac{R_{kd}}{x''_{kdd}} & \frac{R_{kd}}{x''_{kdf}} & \frac{-R_{kd}}{x''_{kd}} & 0 & 0 & 0 \\ -1-\frac{R_{L2}(t)}{x''_d} & \frac{R_{L2}(t)}{x''_{df}} & \frac{R_{L2}(t)}{x''_{kdd}} & \frac{-(R_Q-R_{L1}(t))}{x''_q} & \frac{R_Q-R_{L1}(t)}{x''_{kqq}} & -\sqrt{2}\,\frac{R_{L4}(t)}{x_0} \\ -\sqrt{2}\,\frac{R_{L3}(t)}{x''_d} & \sqrt{2}\,\frac{R_{L3}(t)}{x''_{df}} & \sqrt{2}\,\frac{R_{L3}(t)}{x''_{kdd}} & -\sqrt{2}\,\frac{R_{L4}(t)}{x''_q} & \sqrt{2}\,\frac{R_{L4}(t)}{x''_{kqq}} & -\frac{(R_0+R_{L0})}{x_0} \end{bmatrix}$$

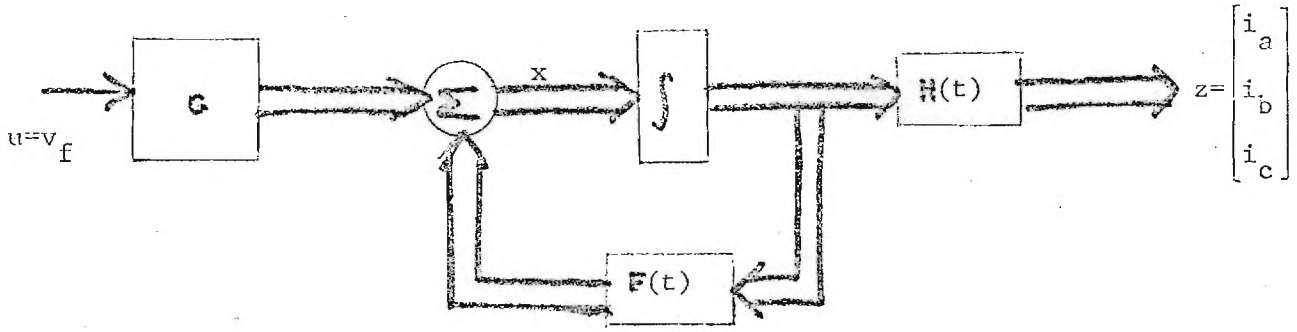
$$G=\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

where $R_D = R_d + R_{L0}$

$R_Q = R_q + R_{L0}$

and $R_{L0} = 1/3(R_a + R_b + R_c)$.

$$H=\sqrt{2/3}\omega\begin{bmatrix} \frac{-\cos\omega t}{x''_d} & \frac{\cos\omega t}{x''_{df}} & \frac{\cos\omega t}{x''_{kdd}} & \frac{\sin\omega t}{x''_q} & \frac{-\sin\omega t}{x''_{kqq}} & -\frac{1}{\sqrt{2}x_0} \\ \frac{-\cos(\omega t-120^\circ)}{x''_d} & \frac{\cos(\omega t-120^\circ)}{x''_{df}} & \frac{\cos(\omega t-120^\circ)}{x''_{kdd}} & \frac{\sin(\omega t-120^\circ)}{x''_q} & \frac{-\sin(\omega t-120^\circ)}{x''_{kqq}} & -\frac{1}{\sqrt{2}x_0} \\ \frac{-\cos(\omega t+120^\circ)}{x''_d} & \frac{\cos(\omega t+120^\circ)}{x''_{df}} & \frac{\cos(\omega t+120^\circ)}{x''_{kdd}} & \frac{\sin(\omega t+120^\circ)}{x''_q} & \frac{-\sin(\omega t+120^\circ)}{x''_{kqq}} & -\frac{1}{\sqrt{2}x_0} \end{bmatrix}$$



Thus F and H are periodically time varying with period equal to $\frac{2\pi}{\omega}$. Notice that for the balanced load $R_a=R_b=R_c=R$, $R_{L1}(t)=R_{L2}(t)=R_{L3}(t)=R_{L4}(t)=0$ and the zero sequence variables $i_0(t)$ and $\lambda_0(f)$ are identically zero, thus $F(t)=F$. In fact, F for the balanced load case is just the previous matrix A with R_d replaced by R_d+R and R_q replaced by R_q+R.

Because the computations will be carried out on a digital computer and the measurements will most likely be in the form of sampled data, the continuous time model given above will be replaced by a discrete-time model described by difference equations. Consider the differential equation

$$\dot{x}(t) = F(t)x(t) + G(t)u(t). \qquad [1]$$

The general solution to this equation is well known (1) to be:

$$x(t) = \Phi(t, t_0)\, x(t_0) + \int_{t_0}^{t} \Phi(t,\tau)G(\tau)u(\tau)d\tau, \qquad [2]$$

where $\Phi(t,\tau)$ satisfies $\frac{d}{dt}\Phi(t,\tau)=F(t)\Phi(t,\tau)$ and $\Phi(\tau,\tau)=I$. $\Phi(t,\tau)$ is called the

state transition matrix. Now consider sampling the continuous-time solution at $t_1, t_2 \ldots, t_k, t_{k+1}, \ldots, t_K$. The solution from t_k to t_{k+1} is (simplifying the notation so that $x(t_k)$ becomes $x(k)$):

$$x(k+1) = \Phi(t_{k+1}, t_k)x(k) + \int_{t_k}^{t_{k+1}} \Phi(t_{k+1},\tau)G(\tau)u(\tau)d\tau.$$

Assuming that $u(t)$ is constant on $t_k \leq \tau < t_{k+1}$, and defining $\Phi(t_{k+1}, t_k)$ and $\Gamma(k) \triangleq \int_{t_k}^{t_{k+1}} \Phi(t_{k+1},\tau)G(\tau)d\tau$, then $x(k+1)=\Phi(k)x(k)+\Gamma(k)u(k)$. If the system is stationary $\Phi(k)=\Phi=e^{FT}$, where $T=t_{k+1}-t_k$, and $\Gamma(k)= = \int_0^T \Phi(v)dvG= \int_0^T e^{FV}dvG=F^{-1}(e^{FT}-I)G$. The matrix exponential $e^{FT} \triangleq \sum_{i=0}^{\infty} \frac{(FT)^i}{i!}$, therefore, an approximate formula, valid for small T, is

$$e^{FT} \cong I + FT.$$

As a result

$$\begin{aligned} x(k+1) &= (I + FT)x(k) + F^{-1}(I + FT-I)G\, u(k) \\ &= (I + FT)\, x(k) + TGu(k). \end{aligned} \qquad [3]$$

The formulation presented above assumes that the system is deterministic, that is no noise is present. To extend the formulation to include consideration errors introduced by noisy systems, assume that additive noise terms corrupt both the input signal u(k) and the measurement z(k).

$$x(k+1) = \Phi(k)x(k) + \Gamma(k)u(k) + \omega(k) \qquad [4]$$

$$z(k) = H(k)x(k) + v(k) \qquad [5]$$

where

x= n-dimension state vector — z= p-dimension observation vector

u= m-dimension input vector — ω= m-dimension input noise vector

and v= p-dimension observation noise vector.

$$E[\omega(k)]= E[v(k)] = E[x(0)]=0, \qquad E[x(0)x^T(0)]= \Psi$$

$$E[\omega(k_1)\omega^T(k_2)] = Q\,\delta(k_1-k_2), \qquad E[v(k_1)v^T(k_2)]=R\delta(k_1-k_2) \qquad [6]$$

where $E[\xi]$ denotes expected value of ξ, and

where $\delta(k_1-k_2)$ is Kronecker's delta defined by

$$\delta(k_1-k_2) = \begin{cases} 1 \text{ if } k_1 = k_2 \\ 0 \text{ otherwise} . \end{cases}$$

Some of the parameters of Φ, Γ, H, R, and Q are not known and the problem is to estimate them. Let the unknown parameters be denoted by the vector α. For the present case α is a constant since all the time-varying terms are already known. Thus this problem is easier than the most general case which would require tracking time-varying parameters. A further simplification of a balanced resistive load yields the even easier case of constant parameters in a stationary system. The most promising approach to the problem is to identify as many parameters as possible from the stationary case (such as a suddenly applied three-phase to ground short circuit on the armature terminals), then to identify the rest of the parameters from the non-stationary system arising from an unsymmetrical load (such as $R_a = 1$ per unit and $R_b = R_c = 0$). It is clear that the zero-sequence parameters cannot be determined from any balanced load tests unless input noise is allowed to drive the zero sequence equations, for in the deterministic system, the zero sequence equation is uncoupled from the rest of the system. Some sort of unsymmetrical load will probably be needed to achieve accurate identification of R_0 and x_0.

III. General Approach to System Identification

The basic approach to modeling the system can be expressed as a three-step process (6):

1. Hypothesize a reasonable model
2. Estimate the parameters of that model
3. Test the validity of the model.

Obviously, if errors are made in step one, the parameter estimation will yield disappointing results. A brief introduction to the model which has been hypothesized was presented in the previous section of this report. The bulk

of the rest of the report will be dedicated to comparison of three methods to do step two. Step three is beyond the scope of this report; however, it is considered to the extent that it bears upon the relative merits of the various estimation schemes.

In particular, one important criterion by which a parameter estimation method will be judged is the ease with which it allows implementing a reasonable test of the model validity. Further, the scheme must be general enough to handle the model set forth, a time-varying linear, discrete-time state variable representation. Of course, the generator itself is a nonlinear device and it is highly desirable that the estimation method be general enough to handle nonlinearities such as saturation when and, indeed, if the model is generalized to include such effects. The proposed model includes a known input v_f which is probably very large compared to the noise $w(k)$, and further, the input is usually derived from a dc source (such as another machine or a rectifier of some sort) which cannot be varied at will. Thus constraints are put upon the types of inputs available, that is the input signal is not as easily manipulated as it is in many system identification problems. The final consideration is perhaps the stickiest point of all; that is, how should the a priori information (or lack of information) about the parameters and the noise statistics be taken into account. This point is discussed in detail later.

To aid in comparing parameter estimation techniques, it is helpful to briefly review some concepts from statistics (2, 6, 7) dealing with the theoretical performance of estimators. The first desirable property of estimators is that they produce unbiased estimates. An unbiased estimate, $\hat{\alpha}$, of the parameter vector α means $E[\hat{\alpha}] = \alpha$ (or if α is considered a random variable, $E[\hat{\alpha}] = E[\alpha]$) where E denotes expected value. A consistent estimator is one whose estimates converge in probability, i.e.

$$\lim_{k \to \infty} p\left\{(\hat{\alpha}_k - \alpha) = 0\right\} = 1.$$

An efficient estimate is an unbiased estimate which has a smaller error covariance = $E[(\alpha-\alpha)^2]$ than any other unbiased estimate. An estimator which is unbiased or efficient only in the limit as $k\to\infty$, is called asymptotically unbiased or asymptotically efficient. All of these criteria are measures of statistical "power" of estimators and provide valuable insight into their theoretical performance.

From the practical point of view, the performance of an estimator depends on computational considerations as well as the nice theoretical properties which it may possess. The most obvious considerations of pratical importance are ease of implementation of the estimator and the speed with which the estimates converge (hopefully) to the correct values. As might be expected, some of these computational considerations can be traded off against theoretical performance. This idea will be pursued in detail in the succeeding section. Judging from the bulk of the literature (survey papers 14, 15, and 17) one of the most important practical problems is that of divergence of the estimator. Even in cases where good theoretical performance is expected, the estimates can diverge from the true value of the parameters. Since this phenomena depends quite heavily on the particular system under analysis, on whether a priori information is accounted for properly, and on the existence of structural errors in the model (e.g. a poor model proposed) this problem is of paramount importance and presents many pitfalls to the unwary.

In the next section, all these factors are considered for three basic identification methods (while hopefully avoiding most pitfalls).

IV. Comparison of Three Identification Schemes

This section is dedicated to the comparison of three different, yet related, methods of system identification by parametric methods. Before

beginning the detailed analysis of each method, an overview of the three is given to clarify their relationships (15: pp. 209-212 , also 3, 4, 5, 6, 24).

Assume first that the parameter vector α is a random variable and that the a priori probability density function of α, i.e. $p(\alpha)$, is known. Further assume that the prior statistics of the noise are known, or equivalently that $p(z|\alpha)$ is known. Define a cost function $c(\hat{\alpha}, \alpha)$ which represents the cost of choosing estimate $\hat{\alpha}$ if α is the true parameter and which has a minimum value at $\hat{\alpha}=\alpha$. The average risk of choosing $\hat{\alpha}$ is

$$R(\hat{\alpha}) = E_{\alpha}\left\{ E_{z}[c(\hat{\alpha}, \alpha)|\alpha]\right\} = \int_{m} \int_{k+1} c(\hat{\alpha}, \alpha)p(z|\alpha)p(\alpha)d^{k+1}z d^{m}\alpha$$

where $\int_{m}$ means an m-fold integral

and $d^{m}\alpha = d\alpha_{1}\, d\alpha_{2} \ldots d\alpha_{m}$

and E_{α} means expectation with respect to α.

Minimizing this risk leads to the so-called Bayesian estimate of α. Applying Bayes rule and interchanging the order of integration leads to

$$R(\hat{\alpha}) = \int_{k+1} d^{k+1}z \left\{ p(z) \int_{m} c(\hat{\alpha}, \alpha)p(\alpha|z)d^{m}\alpha \right\}.$$

Minimization of $R(\hat{\alpha})$ is accomplished by minimizing the second integral since $p(z)>0$, and this is accomplished at the maximum value of $p(\alpha|z)$. Thus the Bayesian estimator is often called the maximum a posteriori estimator (abbreviated MAP estimator). The MAP estimator can also be used even if the cost function is not known.

Next suppose that neither the cost function nor the prior statistics of α are known. Assume then that $p(\alpha) = \lambda$, a constant, thus

$$\max_{\alpha}[p(\alpha|z)] = \frac{\lambda}{p(z)} \times \max_{\alpha}[p(z|\alpha)].$$

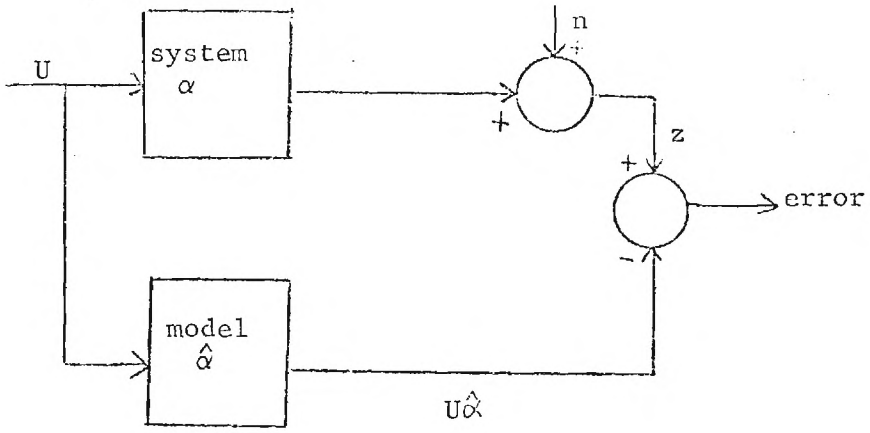
The MAP estimator reduces to one which maximizes the conditional probability of the observed output given the parameter vector α is known. Thus the MAP

estimator "looks like" a maximum likelihood estimator, but being precise the latter is one which maximizes the likelihood function = $p(z:\alpha)$ where x is an unknown parameter vector but is <u>not</u> a random variable (15,6). This is a purely conceptual difference, but it is important since both the MAP estimator and the ML estimator are valid approaches in their own right. Thus neither is really a special case of the other.

It will be shown later that for the case where the noise is normally distributed with mean zero and covariance R, the maximum likelihood estimator is found by minimizing a quadratic cost functional, J. Let the observation be

$z(k) = h(x,k) + n(k)$

n = Gaussian noise, mean zero, covariance R



$$U = \begin{bmatrix} U_1(0) & \cdots & U_m(0) \\ \vdots & & \vdots \\ U_1(k) & & U_m(k) \end{bmatrix}$$

Then (15: p. 211)

$$J = (z-U\hat{\alpha})^T R^{-1} (z-U\hat{\alpha})$$

Maximizing this functional yields

$$\hat{\alpha}_{ML} = (U^T R^{-1} U)^{-1} U^T R^{-1} z.$$

Thus the maximum likelihood estimator is linearly related to the observation for this case. It can also be shown that it is unbiased and efficient (15, 5, 6) for this case. Notice that if any arbitrary positive definite matrix is

is used in place of the inverse covariance kernel of the noise N^{-1}, this approach yields the well-known weighted-least square estimate (3, 4, 5, 6, 15, 26). Thus maximum likelihood estimation is a special case of weighted least squares wherein the weighting matrix is chosen according to the noise statistics.

To allow the detailed comparison of these three methods of parameter estimation, the system given by equations [4], [5], [6] with unknown parameter vector α is formulated in terms of maximum a posteriori, maximum likelihood and weighted least squares estimation schemes.

The maximum a priori state estimator is the well-known Kalman-Bucy filter (4, 19). By augmenting the state vector to include the parameter as well (5, 6) there follows:

$$\tilde{x}(k) = \begin{bmatrix} x(k) \\ \alpha(k) \end{bmatrix}$$

where $\alpha(k+1) = \alpha(k)$

$$\tilde{x}(k+1) = \phi[\tilde{x}(k), u(k), k] + \Gamma[\tilde{x}(k), k]\omega(k)$$

$$z(k) = h[\tilde{x}(k), k] + v(k).$$

$\omega(k)$ and $v(k)$ are independent, zero mean, Gaussian-Markov white noise processes with

$$\text{cov } \omega = E[\omega(k_1)\omega^T(k_2)] = Q(k)\delta(k_1-k_2)$$

$$\text{cov } v = E[v(k_1)v^T(k_2)] = R(k)\delta(k_1-k_2).$$

$x(0)$ is independent of $\omega(k)$ and $v(k)$ and is Gaussian with mean $\mu x(0)$ and variance Ψ. To simplify the notation, let x now be the augmented state vector including the parameters α, and let sequences $\{x(k)\}$ and $\{z(k)\}$, $0 \leq k \leq K$, be denoted by X(K) and Z(K) respectively. By use of Bayes' rule

$$p[X(K)|Z(K)] = \frac{p[Z(K)|X(K)]p[X(K)]}{p[Z(K)]}$$

$$p[Z(K)|X(K)] = \prod_{k=1}^{K} \frac{\exp\{-\frac{1}{2}(z(k)-h[x(k),k])^T R^{-1}(k)(z(k)-h[x(k),k])\}}{(2\pi)^{p/2}\det[R^{-1}(k)]^{\frac{1}{2}}}$$

since with x(k) given, z(k) must be Gaussian because v(k) is Gaussian. Also, using the definition of conditional probability and the x(k) is a Marleov process

$$p[x(0)] \prod_{k=1}^{K} p[x(k)|x(k-1)].$$

$p[x(k)|x(k-1)]$ is Gaussian with mean $\Phi(x(k-1), k-1)$ and variance $[x(k-1), k-1]$ $p[Z(K)]$ is just a scale factor containing no terms involving x(k). Thus, the problem becomes maximizing

$$p[X(K)|Z(K)] = A \exp\{-J\}$$

where

$$J = \tfrac{1}{2}[x(0)-\mu_x(0)]^T \Psi^{-1}[x(0)-\mu_x(0)]$$
$$+ \tfrac{1}{2}\sum_{k=1}^{K} (z(k)-h[x(k),k])^T R(k)^{-1}(z(k)-h[x(k),k])$$
$$+ \tfrac{1}{2}\sum_{k=1}^{K} \omega^T(k-1)Q^{-1}(k-1)\omega(k-1)$$

and where A is not a function of x. So the problem is cast into one involving the minimization of a quadratic cost functional. Unfortunately, it is generally not possible to solve for the values of Q(k) and R(k) and, as a result, some kind of linearization or approximate methods must be used. The literature on this subject is quite rich (see 5 and 6 for example) and the detailed analysis of this problem is not pursued here. The main difficulty with this approach is the great computational effort required (17). Another drawback is that if incorrect prior statistics are used, the results of the Bayesian approach can be worse than those of other, simpler methods.

$$\hat{x}(k+1|k+1:\alpha) = \Phi(k)\hat{x}(k|k:\alpha) + \Gamma(k)u(k)$$
$$+ V(k+1|k:\alpha)H^T(k+1)V_\delta^{-1}(k+1|k:\alpha)\delta(k+1:\alpha),$$

$$\delta(k+1:\alpha) = z(k+1) - H(k+1)\Phi(k)\hat{x}(k|k:\alpha) + \Gamma(k)u(k),$$

$$V(k+1|k:\alpha) = V(k+1|k:\alpha)\,\Phi^T(k) + Q(k)$$

$$V(k+1|k+1:\alpha) = V(k+1|k:\alpha) - V(k+1|k:\alpha)H^T(k+1)V_\delta^{-1}(k+1|k:\alpha)$$
$$- V(k+1|k:\alpha)H^T(k+1)V_\delta^{-1}(k+1|k:\alpha)H(k+1)V(k+1|k:\alpha)$$

$$V_\delta(k+1|k:\alpha) = R(k+1) + H(k+1)V(k+1|k:\alpha)H^T(k+1)$$

$$V(0|0) = \Psi \qquad \hat{x}(0|0) = 0$$

Thus even though the requirement to estimate the state vector was relaxed, we find that it still must be done to implement this method. Notice that this problem is somewhat simpler than the one resulting from the Bayesian-state augmentation approach. The maximum likelihood formulation, then reduces to the minimization of a quadratic cost functional, equation [8], and the implementation of a Kalman filter. Of course slightly different forms of the optimal state estimator such as the smoothing form could have been used here (4, 5, 6). Also, quite a variety of computational approaches to the resulting minimization problem (or the equivalent two point boundary value problem) have been proposed (see for example 5, pp. 152-209).

The third approach to system identification is the much more direct and intuitive approach of weighted least squares--model reference. This technique is basically an extension of the idea of curve-fitting by the method of least squares. The minimization of a quadratic cost functional containing weighting matrices whose elements are chosen simply by "engineering judgement" rather than by sophisticated statistical analyses has great appeal from the computational point of view and is a reasonable engineering solution to the often intractable problems raised by the first two methods. To make these ideas precise, assume that we consider a model of the system,

The preceding discussion considered directly estimating both the state vector and the parameter vector. Now, in an attempt to estimate just the parameters, consider the formulation in terms of maximum likelihood estimation (abbreviated MLE) (5, 6, 20). This approach considers the parameter vector α to be unknown but not a random variable, thus the notation $p(z:\alpha)$ means the probability density of the measurements given the parameters α. This density function is known as the likelihood function. Considering once again the system given by equations [4] and [5], assume that the noise has mean zero and covariances given by equations [6]. Using the same notation as before

$$p[Z(k):\alpha] = p[Z(k-1),\ z(k):\alpha] = p[Z(k-1):\alpha]\,p[z(k)|Z(k-1):\alpha], \qquad [7]$$

by the product rule. Now let

$$\hat{z}(k|k-1:\alpha) = E[z(k)|Z(k-1):\alpha]$$

$$\delta(k:\alpha) = z(k)-\hat{z}(k|k-1:\alpha)$$

and $v_\delta(k|k-1:\alpha) = E[\delta(k:\alpha)\delta^T(k:\alpha)]$.

$\delta(k:\alpha)$ is called the innovations process and represents new information added by the measurement $z(k)$ (4). If $z(k)$ is assumed to be a Gaussian process then

$$p[z(k)|Z(k-1):\alpha] = \frac{\exp[-\frac{1}{2}\,\delta^T(k:\alpha)v_\delta^{-1}(k|k-1:\alpha)\,\delta(k:\alpha)]}{(2\pi)^{p/2}[\det v_\delta(k|k-1:\alpha)]^{\frac{1}{2}}}$$

and the likelihood function is

$$p[Z(K):\alpha] = \prod_{k=1}^{K} p[z(k)|Z(k-1):\alpha].$$

Define the log likelihood function

$$L(K:\alpha) = \ln\ p[Z(K):\alpha].$$

Thus equation [7] becomes

$$2L(K:\alpha) = -KP\ln(2\pi) - \sum_{k=1}^{K} \ln\ \det[v_\delta(k|k-1:\alpha]$$

$$-\sum_{k=1}^{K}[\delta^T(k:\alpha)V_\delta^{-1}(k|k-1:\alpha)\delta(k:\alpha)]. \qquad [8]$$

The equations for v_δ and δ are given by the Kalman filter algorithms, one form of which is summarized below (6). Let $\hat{x}(k|k:\alpha)$ be the estimate of $x(k)$ using measurement $Z(k)$.

$$x_m(k+1) = \Phi(k)x_m(k) + \Gamma(k)u_m(k)$$

$$z_m(k) = H(k)x_m(k),$$

where α=unknown parameters of Φ, Γ, H. Choose α, $u_m(k)$, and $x_m(k)$ to minimize a reasonable cost functional. One such functional is (6)

$$J = \sum_{k=1}^{K} [z(k)-H(k)x_m(k)]^T M[z(k)-H(k)x_m(k)]$$

where M is apositive definite weighting matrix. Other more complicated cost functionals can be proposed (6), but this simple one illustrates the procedure. Notice that the weighted least squares approach is very similar to the cost functionals derived for the maximum likelihood formulation. In fact, MLE can be considered as a special case of weighted least squares (WLS) where the weighting matrix, or matrices, are chosen in a particular way to give good statistical properties (6, 14, 15, 20, 22, 26). Thus we can "trade off" the theoretical performance to gain computational simplicity by using WLS estimation.

The three approaches to parameter estimation briefly discussed above make interesting contrasts with each other. The Bayesian approach, theoretically the best if all prior statistics are known, can be the worst from a practical point of view. Not only is it the most difficult to implement, but it suffers from a tendency to give grossly incorrect estimates if the wrong prior statistics are used (17). Various attempts have been made to reduce this problem (16, 23) but it remains a fundamental practical problem. The maximum likelihood techniques, easier to implement and less computationally expensive (6, 17, 20) nevertheless are theoretically quite powerful. If an efficient estimate exists the maximum likelihood estimator will be efficient. Also, for linear models with Gaussian statistics the ML estimate is consistent and unbiased (6, 7).

The weighted least squares method is, of course, the simplest computationally but can give bad results, e.g. large biases, in cases where the ML estimator works well (6). The table below is a summary of relative merits of the three schemes with regard to ease of implementation, computational expense and the best theoretical accuracy.

	ease of implementation	computational expense	theoretical accuracy
a) Bayesian-State Augmentation	hardest	great	best
b) Maximum Likelihood	somewhat easier than (a)	somewhat less than (a)	almost as good as (a)
c) Weighted Least Squares	easiest	least	worst

"Theoretical accuracy" refers to the best possible accuracy under the condition that any assumptions made during the derivation are valid. As already pointed out, the first two schemes (and particularly the first, according to Sarĭdis in reference 17) often result in divergence problems. One often-used approach to solve this is to fall back on the weighted least squares scheme, where by some "twiddling" of the parameters of the weighting matrices it is often possible to prevent divergence and get reasonably good results. It is, of course, possible to include some prior information in the weighting matrix of the WLS approach by engineering judgement. Another drawback to WLS is that it does not provide as good an indication of how well the model fits as the statistically oriented methods, so once again the models must use his engineering judgement. Any of these techniques could, conceptually at least, be extended to include nonlinear effects like saturation. Practically, however, this extension may well prove to be hard to formulate and impossible to implement (6).

In conclusion, the use of Bayesian estimation on an augmented state-

parameter vector gives rise to a very difficult nonlinear filtering problem which can be solved in many cases by the extended Kalman filter. A more direct formulation which is more efficient computationally is parameter estimation by maximum likelihood techniques. The simplest method discussed, the weighted least squares-model reference, ignores the statistical nature of the problem but can give acceptable results for cases which are not computationally feasible by maximum likelihood or Bayesian approaches.

V. A Numerical Example

As a preliminary investigation of the possibility of identifying machine parameters via system identification techniques, the following simplified case was tested. It was assumed that a balanced load was applied to the machine and that $X''_{kd} \to \infty$. Furthermore, it was assumed that λ_{kd} was small enough to be ignored. With these assumptions the dynamical model of the machine reduces to

$$\dot{\tilde{x}} = \begin{bmatrix} \dot{\lambda}_d \\ \dot{\lambda}_q \\ \dot{\lambda}_f \end{bmatrix} = \begin{bmatrix} -R_D/X''_d & \omega & R_D/X''_{df} \\ -\omega & -R_q/X''_q & 0 \\ R_f/X''_{df} & 0 & -R_f/X''_f \end{bmatrix} \begin{bmatrix} \lambda_d \\ \lambda_q \\ \lambda_f \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} v_f \triangleq \tilde{F}\,\tilde{x} + \tilde{G}\,v_f \qquad [9]$$

A simulation lasting 4 milliseconds was performed using an integration time step of 10 microseconds. A unit step function was used for the field voltage and the initial state was

$$[\lambda_d(0) \quad \lambda_q(0) \quad \lambda_f(0)] = [0 \quad 0 \quad .0937]$$

The values used in the simplified machine model coefficient matrix were

$\omega = 2513$ $\qquad R_D/X''_{df} = 1151$

$R_D/X''_d = 1288$ $\qquad R_f/X''_{df} = 66.95$

$R_q/X''_q = 294.3$ $\qquad R_f/X''_f = 70.51$

The values of the fluxes generated were recorded in a data file and used as an input to a weighted-least-squares-identification program developed by Taylor and Iliff at the NASA Langley Research Center (27).

V.I. The Identification Algorithm

The program uses a modified Newton-Raphson method to solve the optimization problem and generates the Cramer-Rao bound as an estimate of the covariance of the parameter estimates. A brief derivation of the algorithm follows.

The cost function is modified to take into account the initial guess of the parameters and the state.

$$J = \sum_{k=1}^{K} [z(k) - z_m(k)]^T M_1^{-1} [z(k) - z_m(k)]$$

$$+ [\tilde{x}(0) - \hat{x}(0)] P_0^{-1} [\tilde{x}(0) - \hat{x}(0)]$$

$$+ [c - c_0] M_2^{-1} [c - c_0] \qquad [10]$$

where z is the recorded measurement, z_m is the measurement generated by the model, M_1 is either a weighting matrix or the measurement error covariance, $\hat{\underline{x}}(0)$ is the a priori best estimate of the initial state and P_0 is either the error covariance associated with that estimate or a weighting matrix c is a vector whose coefficients are the unknown parameters, c_0 is the initial estimate of c and M_2 is either the error covariance associated with that estimate or a weighting matrix.

We observe that small perturbations in the unknown parameters and the unknown initial state will cause the following small perturbation in z_m

$$z_m(k) = \hat{z}_m(k) + \frac{\partial z_m(k)}{\partial c}\,\delta c + \frac{\partial z_m(k)}{\partial x(0)}\,\delta \underline{x}(0) \qquad [11]$$

where $\hat{z}_m$ are the measurements generated with $c = c_0$ and $\tilde{x}(0) = \hat{x}(0)$, $\delta c = c - c_0$, $\delta x(0) = \tilde{x}(0) - \hat{x}(0)$. Substituting [11] into [10] and completing the squares yields:

$$J = \sum_{k=1}^{K} \delta\hat{z}(k)^T M_1^{-1}\, \delta\hat{z}(k) - \delta\lambda^*\, W\, \delta\lambda^* + [\delta\lambda - \delta\lambda^*]^T W\, [\delta\lambda - \delta\lambda^*] \qquad [12]$$

where

$$\lambda^T = [c^T\ x(0)^T] \qquad \delta\lambda = \lambda - [c_0^T\ \hat{x}(0)^T]^T$$

$$W = \left[P^{-1} + \sum_{k=1}^{K} \frac{\partial z_m(k)}{\partial \lambda}^T M_1^{-1} \frac{\partial z_m(k)}{\partial \lambda} \right]^{-1} \qquad [13]$$

$$P = \begin{bmatrix} M_2 & 0 \\ 0 & P_0 \end{bmatrix}$$

$$\delta z = \hat{z} - \hat{z}_m$$

$$\delta\lambda^* = - W^{-1} \sum_{k=1}^{K} \frac{\partial z_m(k)}{\partial \lambda} M_1^{-1}\, \delta\hat{z}(k) \qquad [14]$$

At this point all of the unknowns have been incorporated into the vector λ and the cost function has been converted into a form in which the first two terms are independent of the unknown parameters and in which the last term is non-negative-definite. Therefore, a value of the parameters which minimizes the cost function is

$$\delta\lambda = \delta\lambda^*$$

which yields the basic recursive formula.

$$\begin{bmatrix} c \\ x(0) \end{bmatrix} = \begin{bmatrix} c_0 \\ \hat{x}(0) \end{bmatrix} + \delta\lambda^* \qquad [15]$$

Observe, that once we have corrected the parameter and initial state estimates via [15], the least squares cost function becomes

$$J = \sum_{k=1}^{K} \delta z(k)^T M_1^{-1} \delta z(k) - \delta\lambda^* W \delta\lambda \qquad [16]$$

The first term on the right side of [16] corresponds to the first term in [10], while the second term represents the amount we have reduced to the cost function. If P(0), M_1 and M_2 are truly covariance matrices, then W is the Cramer-Rao bound on the error covariance matrix associated with our new estimates of c and $\tilde{x}(0)$. Therefore, we repeat the procedure until the uncertainty in the estimates, W, decreases sufficiently. On the other hand, if P(0), M_1 and M_2 are merely weighting matrices, we still repeat the procedure until W becomes sufficiently small--as long as W is large, we can still decrease the cost function.

The partial derivatives needed in [14] are calculated as follows. First, observe that

$$\frac{\partial z_m(k)}{\partial\lambda} = \frac{\partial z_m(k)}{\partial \tilde{x}(k)} \frac{\partial \tilde{x}(k)}{\partial\lambda} \qquad [17]$$

But,

$$\frac{\partial}{\partial c} \dot{\tilde{x}} = \frac{\partial \tilde{F}}{\partial c} \tilde{x} + \tilde{F} \frac{\partial \tilde{x}}{\partial c}$$

or

$$\frac{d}{dt} \left[\frac{\partial \tilde{x}}{\partial c}\right] = \tilde{F} \left[\frac{\partial \tilde{x}}{\partial c}\right] + \frac{\partial \tilde{F}}{\partial c} \tilde{x} \; ; \quad \frac{\partial \tilde{x}(0)}{\partial c} = 0 \qquad [18]$$

where

$$\left[\frac{\partial \tilde{F}}{\partial c}\,\tilde{x}\right]_i \triangleq \sum_{k=1}^{n} \frac{\partial \tilde{F}_{ik}}{\partial c}\,\tilde{x}_i$$

On the other hand, for a given c, we have

$$\tilde{x}(t) = \Phi(t)\,\tilde{x}(0) + \int_0^t \Phi(t-\sigma)\,\tilde{G}(\sigma)\,v_f(\sigma)d\sigma$$

from which we conclude

$$\frac{\partial \tilde{x}(t)}{\partial \tilde{x}(0)} = \Phi(t) \qquad [19]$$

or

$$\frac{d}{dt}\left[\frac{\partial \tilde{x}(t)}{\partial \tilde{x}(0)}\right] = \tilde{F}\left[\frac{\partial \tilde{x}(t)}{\partial \tilde{x}(0)}\right] \;;\quad \frac{\partial \tilde{x}(0)}{\partial \tilde{x}(0)} = I \qquad [20]$$

In summary, the partial derivatives required to correct the values of the unknown parameters and initial conditions are generated by numerically integrating [18] and [20].

One other numerical integration must be performed to generate $\delta\lambda^*$. In [14], the difference between the actual measurements and the measurements generated by the latest model is required. That error is determined by numerically integrating the state equations, [9], with the current best estimate of c and $\tilde{x}(0)$.

V.II Results and Recommendations

The data-file, along with the actual initial value of the state and were fed into the algorithm, and it was asked to generate the actual values of

$$R_D/x''_d \;,\; R_D/x''_{df} \;,\; R_f/x''_{df} \;,\; R_q/x''_q \text{ and } R_f/x''_f$$

The initial guesses of these values were inputed as zeros and the weighting matrix was chosen to be the diagonal matrix which makes the maximum value of each component of the state be unity. After several iterations, the algorithm converged to the values shown in Table I.

TABLE I

Parameter	Actual Value	Initial Estimate	Final Estimate
R_D/x''_d	1288	0	1244.32
R_D/x''_{df}	1151	0	1144.33
R_f/x''_{df}	66.95	0	43.15
R_q/x''_q	294.3	0	296.66
R_f/x''_f	70.51	0	68.01

Plots showing a comparison of the actual step response of the machine, the response predicted by the initial model of the machine and the response predicted by the final model of the machine are given in Figures 1, 2 and 3. An examination of these results shows that the final estimates of R_D/x''_{df} and R_q/x''_q are within 0.4% of their actual values, while the errors in the field circuit parameters (R_f/x''_{df} and R_f/x''_f) are ~~35~~ 3.5% and 2.5%. A closer examination of the simplified model makes the reason for the discrepancy in the field parameters (in particular R_f/x''_{df}) obvious, and shows an easy way to improve the accuracy of R_f/x''_{df}.

A simulation diagram of the simplified model is shown in Figure 4. Observe that the simulation diagram has been partitioned into a slow and fast portion, where the term R_f/x''_f is the reciprocal of the slow partition's (mode's) time constant (14.2 msecs.) and the term R_f/x''_{df} is just a coupling from the fast mode to the slow mode. The fast mode's eigenvalues are $(-791.15 \pm j2463.2)$ which corresponds to a time constant of (1.26 msecs). Finally, observe that the d. c. gain of the slow mode (field circuit) is .0142. From these observations, it is obvious that during

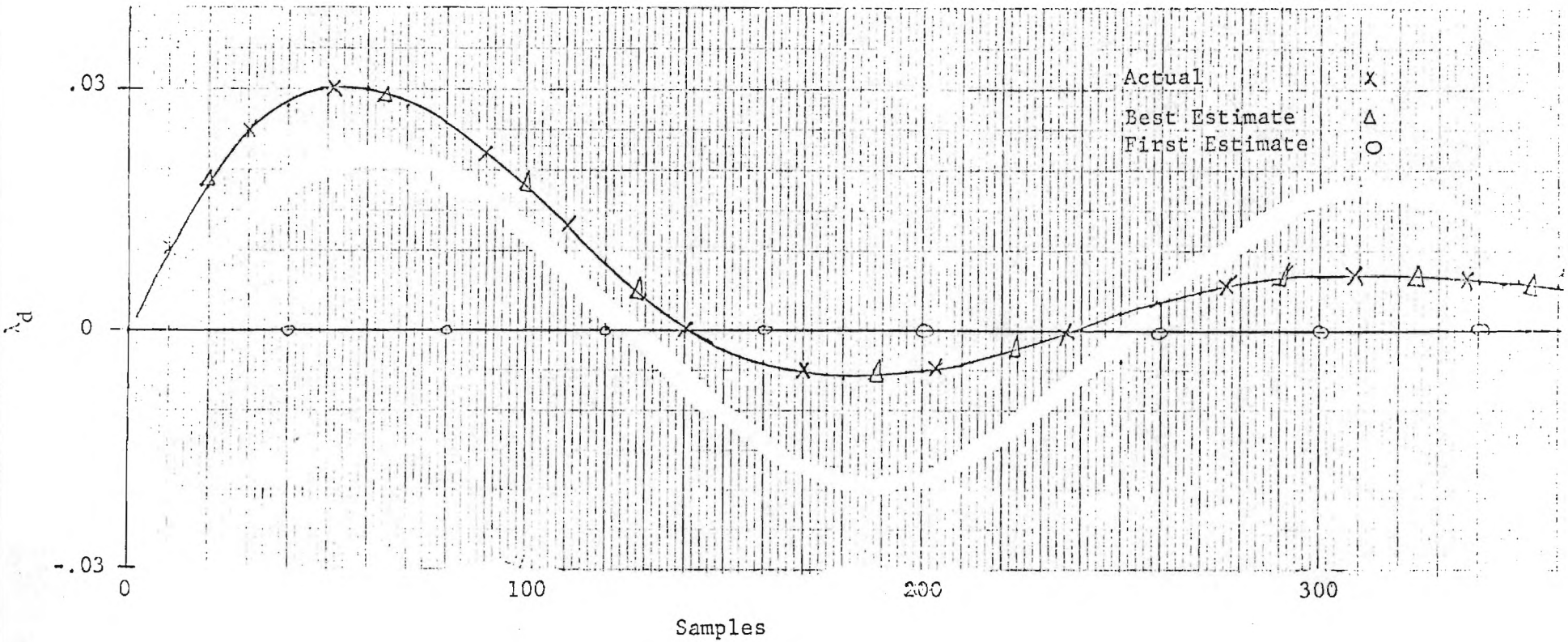


Figure 1 A plot of λ_d versus time.

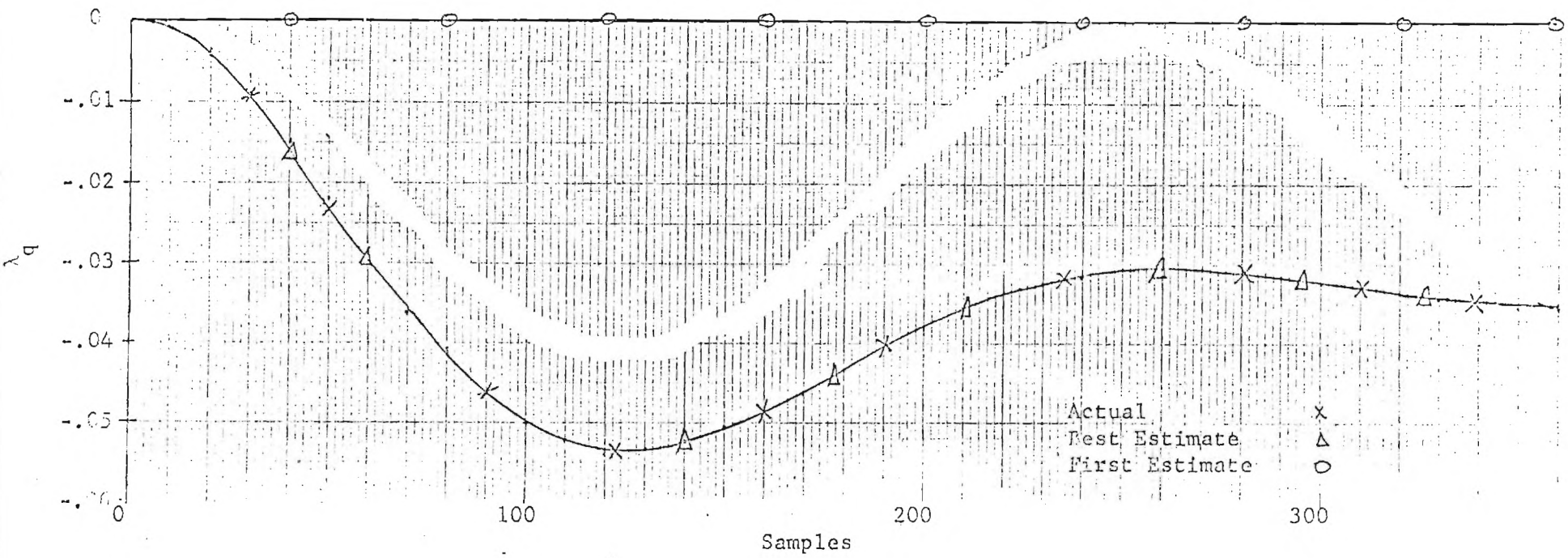


Figure 2 A plot of λ_q versus time

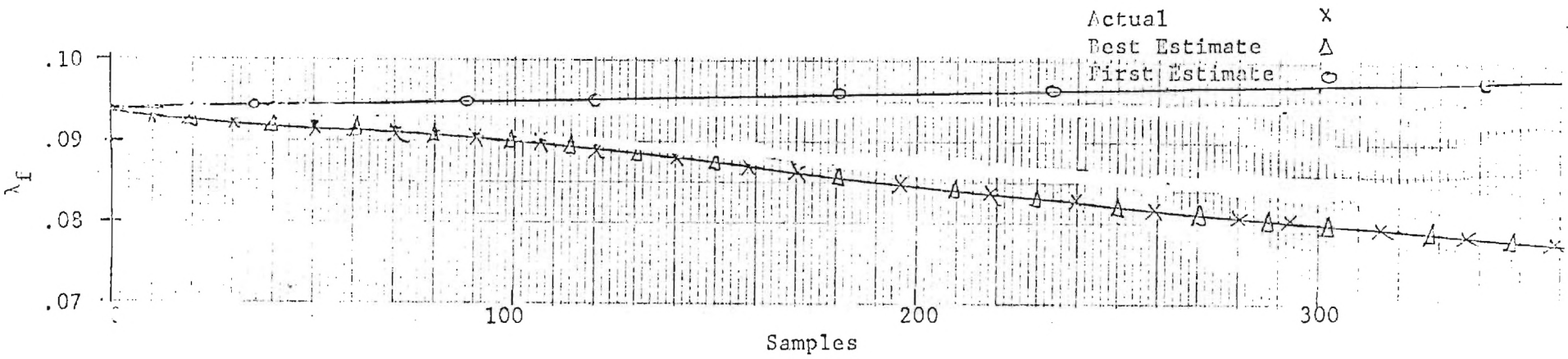


Figure 3 A plot of λ_f versus time.

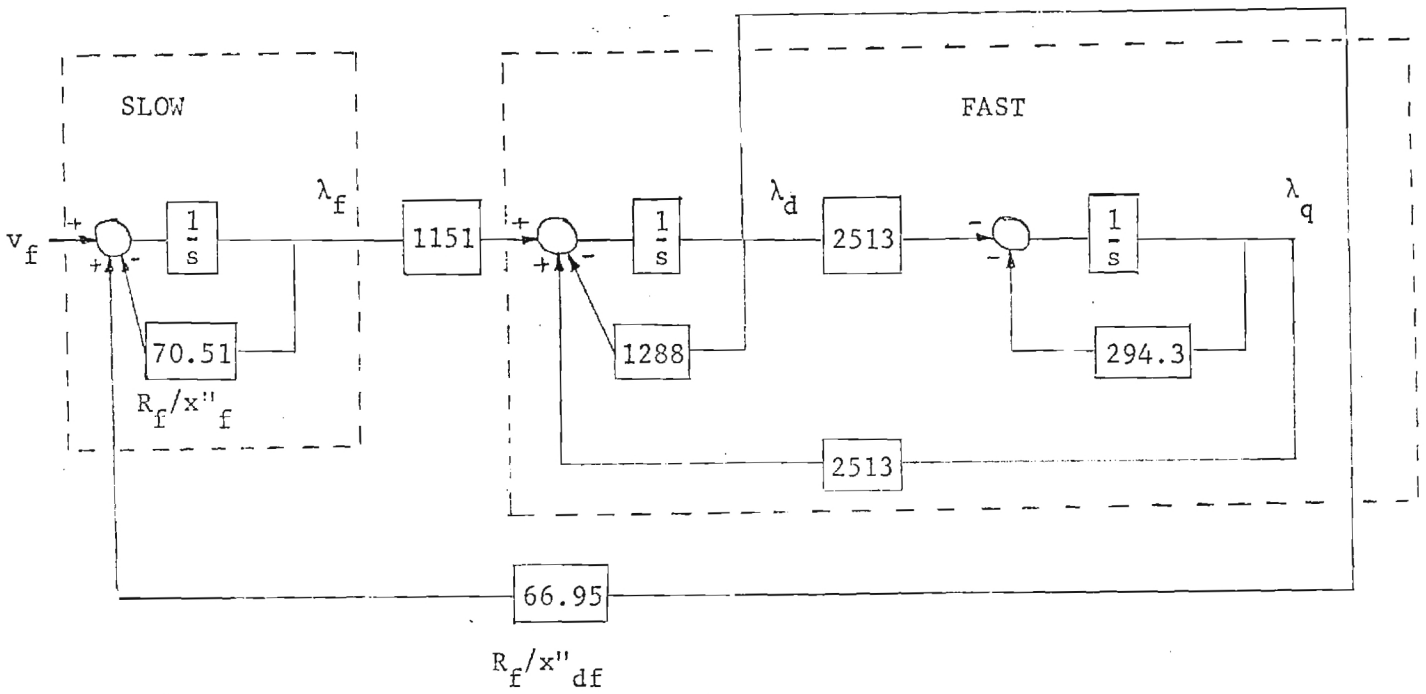


Figure 4 A simulation diagram of the simplified model.

the 4 millisecond simulation, the slow mode (the field circuit) does not have enought time to respond (its time constant is 14.2 milliseconds). This accounts for the 3.5% error in R_f/x''_f and indicates that this error could easily be decreased by increasing the simulation time. The large error in the coupling between the fast mode and the slow field circuit is a result of two things. It is caused by the low gain of the field circuit (.0142) coupled with the fact that, in 4 milliseconds, the slow field circuit does not have time to respond to an input. The result is that, during the first 4 milliseconds, the initial value of λ_f decaying to zero completely dominates the response of the field circuit to any input; therefore, during this time the recorded outputs of the simulation $[\lambda_f \ \lambda_d \ \lambda_g]$ are almost insensitive to R_f/x''_{df}. Since the identification technique used simply "twiddles" the parameters to make the output of the model agree with the recorded outputs, there is no way it can identify a parameter which does not affect the outputs. The solution is conceptually simple, we increase the simulation time to at least four times the slow time constant (4 x 14.2 msec = 56.8 msec.). This will allow the effects of the initial value of λ_f to decay to zero and $\lambda_f(t)$ to become dominated by the input to the field circuit

$$v_f - (R_F/x''_{df})\lambda_d$$

which, in turn, will allow the identification algorithm to identify (R_F/x''_{df}).

Although it is conceptually possible to lengthen the simulation time and complete the identification of the parameters, there is an easier way--break the problem into two parts, the identification of the parameters of the fast mode and the identification of the parameters of the slow mode. The identification of the parameters of the fast mode will be accomplished

using a fast sampling rate and short time duration--ΔT = 10 μsecs. and $t_f - t_0$ = .42 msecs.--, and the machine model in Figure 5. Next we will use a larger sampling period and longer time duration (5 milliseconds and .5 seconds) to identify the slow mode (field circuit). For this relatively large sampling time, we will assume that the fast mode is always in steady state, which corresponds to $\lambda_d = \lambda_q = 0$. From either the simulation diagrams or the dynamical equations we determine the slow-time steady-state model of the fast mode.

$$\begin{bmatrix} \dot{\lambda}_q \\ \dot{\lambda}_d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -R_d/x''_d & \omega \\ -\omega & -R_q/x''_q \end{bmatrix} \begin{bmatrix} \lambda_d \\ \lambda_q \end{bmatrix} + \begin{bmatrix} R_D/x''_{df} \\ 0 \end{bmatrix} \lambda_f$$

or

$$\lambda_d = \frac{R_q}{\omega x''_q} \frac{R_D}{x''_{df}} \left[\omega + \frac{R_D R_q}{\omega x''_q x''_d} \right]^{-1} \lambda_f \triangleq \tilde{K} \lambda_f$$

where ω is known and the fast mode parameters have been identified--$\tilde{K}$ is assumed known. The resulting slow-time model of the machine is given in Figure 6.

Observe that no long simulations with small time steps are required with the proposed slow-fast identification procedure. When the small time step is employed, only a short time duration is required, and, for the long time duration part, a large time step is employed. The suggested procedure is presently being employed to solve the simplified machine model, and the results obtained will be subsequently presented.

VI Summary

Three system identification techniques have been investigated that can be used to identify the unknown parameters in a model of a synchronous electrical generator under severely unbalanced loads. As indicated each

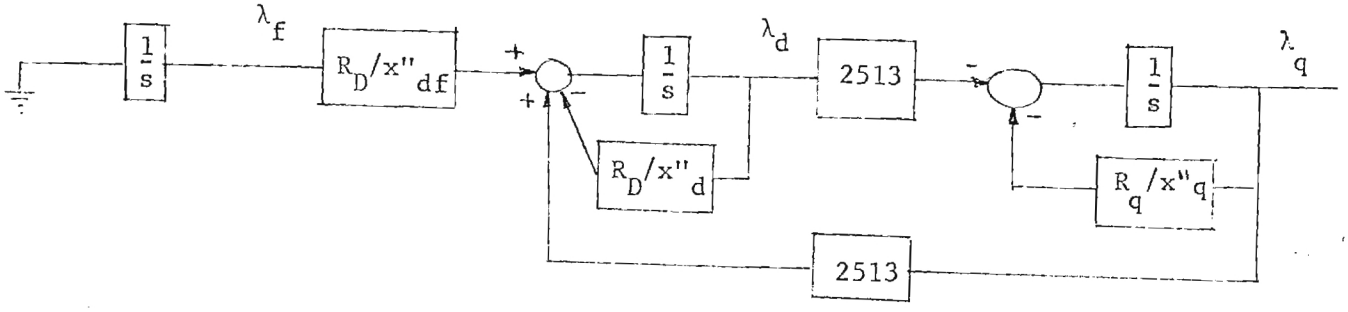


Figure 5 The fast time model.

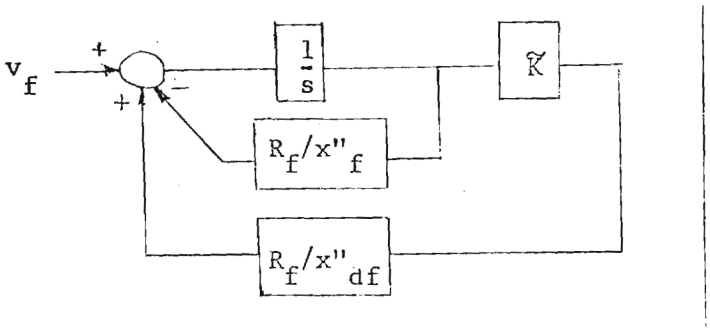


Figure 6 The slow time model.

of the three techniques has advantages and drawbacks; however, all of the techniques can be used. At present the least-square technique has been employed to successfully identify the parameters for a simplified machine model. Finally, a technique has been proposed for alleviating the numerical problems associated with identifying the parameters of a system model which has widely varying time constants.

REFERENCES

Texts

1. Brockett, R. W., Finite Dimensional Linear System, John Wiley and Sons, New York, 1970.

2. Cramer, H., Mathematical Methods of Statistics, Princeton University Press, Princeton, N.J., 1946.

3. Eykhoff, P., System Identification: Parameter and State Estimation, John Wiley and Sons, London, 1974.

4. Sage, A. P. and J. L. Melsa, Estimation Theory with Applications to Communications and Control, McGraw-Hill, Inc., New York, 1971.

5. Sage, A. P. and J. L. Melsa, System Identification, Academic Press,N.Y.,1971.

6. Schweppe, F. C., Uncertain Dynamic Systems, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1973.

7. Wilks, S., Mathematical Statistics, John Wiley and Sons, New York, 1962.

Papers on Machinery

8. Debs, A. S., R. P. Webb, and C. W. Brice, "Summary Report: Analysis of Computational Requirements for Unbalanced Synchronous Generator Simulation," AFAPL Report, November 1974.

9. Debs, A. S., R. P. Webb, and C. W. Brice, "Summary Report: Evaluation of Parameter Accuracy Utilizing IEEE Standards Procedures," AFAPL Report, January 1975.

10. IEEE Standard No. 115, "Test Procedures for Synchronous Machines," March 1965.

11. Park, R. H., "Two-Reaction Theory of Synchronous Machines-Part I," AIEE Transactions, vol. 48, pp. 716-730, July 1929.

12. Rankin, A. W., "Per-Unit Impedances of Synchronous Machine," AIEE Transactions, vol. 64, pp. 569-573, August 1945.

13. Riaz, M., "Hybrid-Parameter Models of Synchronous Machines," IEEE Transactions on Power Apparatus and Systems, vol. PAS-93, No. 3, pp. 849-858, May/June 1974.

Papers On System Identification

14. Astrom, K. J. and P. Eykhoff, "System Identification--A Survey," Automatica, vol. 7, pp.123-162, 1971.

15. Eykhoff, P., "Process Parameter and State Estimation," Automatica, vol. 4, pp. 205-233, 1968.

16. Mehra, R. K., "Approaches to Adaptive Filtering," IEEE Transactions on Automatic Control, vol. AC-17, pp. 693-698, October, 1972.

17. Saridis, G. N., "Comparison of Five Popular Identification Algorithms: A Survey," CDC Paper no. WA3-4, pp. 42-44.

18. Gupta, N. K. and R. K. Mehra, "Computational Aspects of Maximum Likelihood Estimation and Reduction in Sensitivity Function Calculations," IEEE Transactions on Automatic Control, vol. AC-19, No. 6, pp. 774-783, December 1974.

19. Kalman, R. and R. Bucy, "New Results in Filtering and Prediction Theory," ASME Journal of Basic Eng., vol. 83, pp. 95o108, March 1961.

20. Kashyap, R. L., "Maximum Likelihood Identification of Stochastic Linear Systems," IEEE Transactions on Automatic Control, vol. AC-15, No. 1, pp. 25-34, February 1970.

21. Luders, G. and K. S. Narendra, "Stable Adaptive Schemes for State Estimation and Identification of Linear Systems," IEEE Transactions on Automatic Control, vol. AC-19, No. 6, pp. 841-847, December, 1974.

22. Mendel, J. M. "Gradient Estimation Algorithms for Equation Error Formulations," IEEE Transactions on Automatic Control, vol. AC-19, No. 6, pp. 820-824, December 1974.

23. Nabi, N.E. and B. M. Schaefer, "Decision-Directed Adaptive Recursive Estimators: Divergence Prevention," IEEE Transactions on Automatic Control, vol. AC-17, No. 1, pp. 61-68, February 1972.

24. Rhodes, I. B., "A Tutorial Introduction to Estimation and Filtering," IEEE Transactions on Automatic Control, vol. AC-16, No. 6, pp. 688-706, December 1971.

25. Sage, A. P. and G. W. Husa, "Adaptive Filtering with Unknown Prior Statistics," Proceedings of 1969 of 1969 JACC, pp. 760-769.

26. Swerling, P., "Modern State Estimation from the Viewpoint of the Method of Least Squares," IEEE Transactions on Automatic Control, vol. AC-16, No. 6, pp. 707-719, December 1974.

27. Taylor, L. W. and K. W. Iliff, "Systems Identification Using A Modified Newton-Raphson Method-A Fortran Program," NASA Technical Note TN D-6374, May 1972